THE
ECONOMIC
JOURNAL

The Economic Journal, **120** (*March*), 80–104. doi: 10.1111/j.1468-0297.2009.02303.x. © The Author(s). Journal compilation © Royal Economic Society 2009. Published by Blackwell Publishing, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.

CHOICE UNDER UNCERTAINTY: EVIDENCE FROM ETHIOPIA, INDIA AND UGANDA*

Glenn W. Harrison, Steven J. Humphrey and Arjan Verschoor

We review experimental evidence collected from risky choice experiments using poor subjects in Ethiopia, India and Uganda. Using these data we estimate that just over 50% of our sample behaves in accordance with expected utility theory and that the rest subjectively weight probability according to prospect theory. Our results show that inferences about risk aversion are robust to whichever model we adopt when we estimate each model separately. However, when we allow both models to explain portions of the data simultaneously, we infer risk aversion for subjects behaving according to expected utility theory and risk-seeking behaviour for subjects behaving according to prospect theory.

How do individuals in developing countries make choices under uncertainty? Economic theory now provides a rich array of theories to explain this type of behaviour. To answer this question we evaluate two competing theories using data collected from artefactual field experiments[1] conducted in three developing countries (India, Ethiopia and Uganda). Our data consist of real choices from 531 subjects in very poor locales, along with information on individual demographic characteristics.

Our primary objectives are to assess the weight of evidence for the two major received theories of choice under uncertainty, expected utility theory (EUT) and prospect theory (PT), and to assess whether they lead to different inferences about the risk attitudes of our subjects. The importance for development policy of characterising choice behaviour, and hence risk attitudes, is well established. Welfare evaluation of any proposed policy with uncertain outcomes should take into account the aversion that some individuals may have to risk, and the manner in which risk-coping strategies mitigate exposure to risks (Fafchamps, 2004).

Furthermore, it is clear that producers and consumers in developing countries face extraordinarily risky environments in general (Collier and Gunning 1999). The rural poor that make up our subject pool are no different. Fafchamps (2004; p.196) concludes a book-length treatment as follows:

> We have learned that risk affects the rural poor in numerous and profound ways. The magnitude and range of shocks that affect rural populations of the Third World is without comparison in developed economies. Perhaps the only way to describe it to people who have never been there is to compare it to a war economy: death strikes at random a large proportion of the population,

* We are grateful to the UK Department for International Development (award no. R7617) and to the US National Science Foundation (NSF/HSD 0527675 and NSF/SES 0616746) for financial support. Supporting data and statistical code are stored in the ExLab Digital Library at http://exlab.bus.ucf.edu. We are grateful to Robin Cubitt, Oliver Morrissey and two referees for comments.

[1] Our experiments are 'artefactual field experiments' in the terminology of Harrison and List (2004). That is, they involve taking procedures from the laboratory and applying them in the field. The use of field experiments in developing countries has grown dramatically in recent years and is reviewed by Cardenas and Carpenter (2008) and Duflo (2006).

especially children; the provision of health services is either non-existent or insufficient; trade with the rest of the world is difficult so that many commodities are rationed or unavailable and local prices are erratic; food is at times very scarce; and steady wage employment is non-existent so that people must make a living from self-employment in little jobs. To deal with such a harsh environment, people are equipped with very little in terms of advanced technology and accumulated assets. Financial institutions are either absent or inefficient and expensive, and in many places, inflation is rife so that the cost of hoarding money is high.

Thus it is a high priority to obtain accurate characterisations of the risk attitudes of the rural poor. To do so, as we will show, one must also obtain an accurate characterisation of the manner in which choices under uncertainty are made.

Our analysis builds on an experimental tradition that uses field experiments focused on behaviour towards risk, started in India by Binswanger (1980, 1981, 1982). Later contributions include experiments in Zimbabwe by Barr and Genicot (2008), in Chile and Peru by Barr and Packard (2003, 2005), in India, Ethiopia and Uganda by Humphrey and Verschoor (2004*a,b*), and in Timor-Leste by Botelho *et al.* (2005). Humphrey and Verschoor (2004*a,b*) conclude that the behaviour they observed is inconsistent with expected utility maximisation and exhibits subjective probability weighting. They recommend that models of choice under uncertainty in developing countries should replace EUT with a version of PT. This model could then be used in conjunction with the experimental data to evaluate and quantify specific features of behaviour such as attitudes towards risk.

This conclusion echoes calls made on the basis of data collected from numerous experiments conducted in the developed world (Camerer, 1998). However, it is founded on a questionable premise: that the research agenda should establish the single best account of behaviour, and that 'best' should be defined in terms of the single model that explains the data most accurately. What if some subjects are best characterised by one model of choice under uncertainty and other subjects are best characterised by another model of choice under uncertainty, and the two models imply different risk attitudes? It is clear that such a scenario makes it more difficult to inform policy interventions than when one assumes just one model of choice.

To investigate this possibility, and contrary to the approach adopted in conventional studies of risky choice, we take two of the major competing models of risky choice in the literature and allow the data to determine the fraction of behaviour described by each model. Using a 'finite mixture model' approach, we estimate the parameters of each competing model and contrast the results with those emerging from conventional estimates that assume either EUT or PT describe behaviour but not both.

We conclude that there is, in fact, support for each model in our data, so that there is no single, correct model that explains all of the data. Furthermore, as conjectured above, we show that the inferences about parameters of each model differ when one estimates the flexible specification that allows the data to determine the fraction of the choices explained by each model. In particular, under the flexible specification our data point to a concave (risk averse) utility of income function if you assume EUT but a convex (risk-seeking) utility of income function if you assume PT. These results are

© The Author(s). Journal compilation © Royal Economic Society 2009

consistent with the views of Humphrey (2000) and Harrison and Rutström (forthcoming), who argue that it may be inappropriate to search for a single model of risky choice because behaviour is sufficiently heterogeneous that it cannot be described by a single theory. Moreover, this is not the sort of heterogeneity that one can assume to be correlated with observable characteristics of the individuals, although the statistical approach we employ does allow for that.

We review the design of the experimental tasks in Section 1. Each subject made eight choices between pairs of lotteries with real monetary consequences, with outcomes that were substantial in terms of their income and wealth. Each subject was given these experimental tasks as part of a larger household survey, so we also have a set of characteristics to describe the individual and their household. In Section 2 we specify statistical models for these data which allow choices to be made consistently with EUT or PT. We consider all binary choices jointly in order to characterise the decision processes used by our subjects across a range of tasks. The parameters of each theory are allowed to be linear functions of observed individual characteristics, as well as experimental treatments and locations, so we do not assume that every subject has the same utility function or probability weighting function. In Section 2 we also specify a finite mixture model in which observed choices can may be generated by either EUT decision-makers or PT decision-makers. Section 3 presents our results and Section 4 concludes.

1. Experiments

The experimental design provided subjects with an array of choices over monetary lotteries, just as one finds in traditional laboratory experiments in developed countries. Table 1 summarises the design and parameters; an Appendix provides details on procedures and the complete design. All values here are in terms of US dollars and are expressed in cents. We discuss the local currency equivalents, and scaling for purchasing power parity, below.

These particular lotteries were chosen with two considerations in mind. First, we wanted to have pairs of lotteries that presented subjects with several of the classical tests of EUT. These are the tests that have been used in developed country lab experiments

Table 1

Experimental Design

		Lottery A			Lottery B		
Task Type	Task Number	Outcome 1	Outcome 2	Outcome 3	Outcome 1	Outcome 2	Outcome 3
Common Consequence Effect	1	250; 1/4	0; 1/4	100; 1/2	100; 1		
	2	250; 1/4	0; 3/4		100; 1/2	0; 1/2	
	3	250; 3/4	0; 1/4		250; 1/2	100; 1/2	
Cyclical Choice	4	550; 1/2	0; 1/2		250; 3/4	0; 1/4	
	5	250; 3/4	0; 1/4		250; 1/2	200; 1/4	0; 1/4
	6	550; 1/2	0; 1/2		250; 1/2	200; 1/4	0; 1/4
Preference Reversal	7	500; 1/4	0; 3/4		150; 3/4	0; 1/4	
Repeat (one of three possible tasks)	8	250; 3/4	0; 1/4		250; 1/2	100; 1/2	
		250; 3/4	0; 1/4		250; 1/2	200; 1/4	0; 1/4
		500; 1/4	0; 3/4		150; 3/4	0; 1/4	

© The Author(s). Journal compilation © Royal Economic Society 2009

to evaluate EUT (Camerer, 1995), and we wanted to have comparable tasks in developing countries. These tests are carefully calibrated to provide fertile grounds for violation of EUT. Some might argue that they are 'booby trap' or 'trip-wire' tasks, that are so finely calibrated as to make it impossible for EUT to account for behaviour. This is incorrect: if one drew lotteries at random, one would find in many cases that EUT and PT predict exactly the same choice. So all we have done is use our experimental control over the choice of lotteries to focus attention on the domain of (paired) tasks where the data can best discriminate between them.

The second consideration underlying these lottery choices is that we were concerned with literacy, since some of our subjects were bound to be illiterate. We also expected, correctly, that they had no experience with experiments and we wanted a design that we could be as confident as possible that they would understand. For this reason we chose 'salient probabilities' of 0, 1/4, 1/2, 3/4 and 1. This made it easier to explain the operationalisation of risk to the subject, and thereby enhanced credibility. Each outcome was a ball of a particular colour, so a probability of 1/2 would have been, say, 2 green and 2 red balls in a bag and a probability of 1/4 would have been 1 green and 3 red, and so on.

Consider the implementation of task 1 from Table 1 in Uganda as a specific example. Lottery *A* was represented as a red bag containing four coloured marbles. The experiment organiser (Verschoor) placed one yellow marble into the bag and explained that, should this bag be selected and the yellow marble subsequently drawn, it would be worth the equivalent in Ugandan Shillings (Ush) of USD 2.50. This is shown as Outcome 1. Similarly, two green marbles (each worth the local equivalent of USD 1) and one blue marble (worth nothing) were added to the red bag. Lottery *B* was represented by a blue bag of four green marbles, each worth the local equivalent of USD 1. Subjects were provided with a corresponding piece of paper, which showed the contents of the red and blue bags, with appropriate values attached to each differently coloured marble. The subject's task was to indicate which bag of marbles they preferred.

The eight tasks were all *binary* choice lotteries. In each task the subject picked either lottery *A* or lottery *B*. At the end of the experiment one of the eight tasks was selected at random for each subject and the lottery chosen in that task was played-out for real money. This procedure motivates subjects to consider each choice carefully as if it were for real money, rewards them for participation in the experiment and controls for wealth effects. To control for possible order effects, roughly one half of the subjects had the tasks presented in one order and the other half in reverse order.

There were 531 subjects in all. In India there were 223 subjects, drawn from two villages (108 in Vepur and 115 in Guddi). In Uganda there were 208 subjects, again drawn from two villages (107 in Sironko and 101 in Bufumbo). In Ethiopia there were 103 subjects. Mosley and Verschoor (2005) provide more details on these regions and descriptive characteristics of the sample. We have a total of 4,248 actual choices, allowing for some missing responses. With minor variations, the task and procedures were identical across each sample.

In all cases the lotteries were presented in terms of local currency which approximately matched the values shown for the outcomes in Table 1. Our statistical analysis converts the local currency units into US dollars and cents using purchasing power parity (PPP) conversion rates. Thus the statistical analysis is undertaken using our best estimate of the local purchasing power of each monetary outcome. In 2000 the PPP

© The Author(s). Journal compilation © Royal Economic Society 2009

rates for Ethiopia, India and Uganda were 8.22, 44.94 and 1644.47, in terms of the rate at which the local currency converted to one US dollar (Heston *et al.* 2002). The PPP exchange rates actually used for our experiments are close to these: 8.75, 50.33 and 1750, respectively, and differ due to differences in exchange rates prevailing at the exact time of each experiment.[2] We recognise that there can be significant differences in purchasing power within regions of developing countries, reflecting differences in patterns of consumption and local prices (Deaton, 1997; §5.2). Although definitions of poverty differ, there can nonetheless be no doubt that a large fraction of our subjects were closer to absolute poverty lines than conventionally encountered in experiments of this type. The outcomes in our experiment also represented substantial amounts of money to our subjects. Humphrey and Verschoor (2004*b*; p.422) note that the payoffs were 250%, 339% and 278% of prevailing daily wages in Uganda, India and Ethiopia, respectively; the experiments lasted roughly 3 hours for each individual.

2. Alternative Theories

We assume just two competing theories of choice under uncertainty to explain these data: EUT and PT. There are several major alternative theories and many parametric variants of these theories but we take these two theories to be major competitors in the literature.[3] We adopt relatively flexible functional forms to implement each theory.

One of the proposed models is a simple EUT specification which assumes a constant relative risk aversion (CRRA) utility function defined over the 'final monetary prize' that the subject would receive if the lottery were played out. That is, the argument of the utility function is the prize in the lottery, which is always non-negative.[4]

The other model is a popular specification of prospect theory (PT) due to Kahneman and Tversky (1979), in which the utility function is defined over gains and losses separately and a probability weighting function converts the underlying probabilities of the lottery into subjective probabilities. The three critical features of the PT model are

(*i*) that the arguments of the utility function be gains or losses relative to some reference point, taken here to be zero;
(*ii*) that losses loom larger than gains in the utility function; and
(*iii*) that there be a nonlinearity in the transformed probabilities.

[2] To use the specific example from task 1 from Uganda, the non-zero prizes were Ush 5,000 and Ush 2,000, which translated into 286 US cents and 114 US cents, respectively. This is as close as we could come to 250 cents and 100 cents with prize values that were rounded in local currency units. The validity of these lotteries as tests of EUT does not depend on the local currency units exactly matching the values shown in Table 1 but rather on the local currency equivalent of 250 cents being the same across tasks for the same subject.

[3] Starmer (2000) provides an excellent review of the major alternatives. He concludes that if EUT is to be replaced as the dominant theory of risky choice in economics, the evidence points to Tversky and Kahneman's (1992) PT as being the best candidate. Although we reject the notion of completely replacing EUT, as explained later, we accept his view that PT should be viewed as the strongest contender. Since we restrict attention to the gain domain, and typically only have two final outcomes in our lotteries, our implementation of PT is virtually identical to the rank-dependent utility model of Quiggin (1982).

[4] Some take the view that EUT requires that utility be defined over terminal wealth and not income. This is false, as discussed by Cox and Sadiraj (2006) and Rubinstein (2002).

© The Author(s). Journal compilation © Royal Economic Society 2009

The first and second points are irrelevant here since all lottery choices were in the gain domain.[5]

A more complete test of EUT and PT would include loss frames but we did not want to complicate the design of our experiments in the field. Furthermore, many of the early, classic tests of EUT against PT conducted by experimental economists refer solely to the gain domain: for example, see Camerer (1989), Starmer (1992), Harless and Camerer (1994) and Hey and Orme (1994). Thus our differentiation between EUT and PT focuses solely on the properties of the nonlinear probability weighting function, which we discuss more formally below.

Since we do not consider losses, EUT is a special case of PT in which there is no probability weighting. Hence we could just estimate a PT model and test the EUT parametric restriction. This is true, and relevant if one wanted to characterise the data in terms of only one model of choice behaviour. However, our interest is precisely in allowing the observed choice data to be generated by two latent models of choice behaviour and where we do not know *a priori* that observable individual demographic characteristics will allow us to differentiate EUT and PT decision-makers crisply.

2.1. *Expected Utility Specification*

We assume that utility of income is defined by $U(x) = (x^{1-r})/(1-r)$ where x is the lottery prize and r is a parameter to be estimated. With this CRRA specification, $r = 0$ indicates risk neutrality, $r > 0$ indicates risk aversion and $r < 0$ indicates risk loving. Probabilities for each outcome k, $p(k)$, are those that are induced by the experimenter, so expected utility (EU) is simply the probability weighted utility of each outcome in each lottery. Since there were up to 3 implicit outcomes in each lottery i, $\mathrm{EU}_i = \sum_k [p(k) \times U(k)]$ for $k = 1, 2, 3$.

A simple stochastic specification is used to specify likelihoods conditional on the model. The EU for each lottery pair is calculated for a candidate estimate of r and the difference $\nabla\mathrm{EU} = \mathrm{EU}_R - \mathrm{EU}_L$ calculated, where EU_L is the EU of the left lottery in the display and EU_R is the EU of the right lottery. A stochastic choice EUT can then be specified by assuming some cumulative probability distribution function, $\Psi(\cdot)$, such as the logistic.[6] Thus the likelihood, conditional on the EUT model being true, depends on the estimates of r given the above specification and the observed choices. The conditional log-likelihood is

[5] An extension would be to consider parametric reference points that differ from zero and evaluate the structural model implied. This approach is illustrated by Harrison and Rutström (2008; §3.2) in the context of laboratory experiments framed as gains and losses relative to some exogenous endowment. They demonstrate that assuming that subjects use a positive reference point, rather than the one presented in the frame presented by the experimenter, can significantly affect estimates of core structural parameters, such as the extent of loss aversion.

[6] One could extend this analysis to include a behavioural specification of errors conditional on each theoretical model. There are several alternative specifications: see Harless and Camerer (1994), Hey and Orme (1994) and Loomes and Sugden (1995) for the first wave of empirical studies including some formal stochastic specification in the version of EUT tested. There are several species of 'errors' in use, reviewed by Loomes and Sugden (1995). Some place the error at the final choice between one lottery or the other after the subject has decided deterministically which one has the higher expected utility; some place the error earlier, on the comparison of preferences leading to the choice; and some place the error even earlier, on the determination of the expected utility of each lottery.

© The Author(s). Journal compilation © Royal Economic Society 2009

$$\ln L^{\mathrm{EUT}}(r;y,\mathbf{X}) = \sum_i l_i^{\mathrm{EUT}} = \sum_i [(\ln \Psi(\nabla \mathrm{EU}) \times \mathbf{I}(y_i = 1)) + (\ln(1 - \Psi(\nabla \mathrm{EU})) \times \mathbf{I}(y_i = 0))]$$

where $\mathbf{I}(\cdot)$ is the indicator function, $y_i = 1(0)$ denotes the choice of the right (left) lottery in task i and $\mathbf{X}$ is a vector of individual characteristics.

We allow each parameter to be a linear function of the observed individual characteristics of the subject. This is the **X** vector referred to above. We consider six characteristics. Four are binary variables to identify the order of the task, the country, females, and subjects that reported having some secondary education or more. We also included age in years and the number of people living in the household. The estimates of *each* parameter in the above likelihood function entails estimation of the coefficients of a linear function of these characteristics. So the estimate of r, $\hat{r}$, would actually be

$$\begin{aligned}\hat{r} = \hat{r}_0 &+ (\hat{r}_{ORDER} \times ORDER) + (\hat{r}_{ETHIOPIA} \times ETHIOPIA) + (\hat{r}_{UGANDA} \times UGANDA)\\ &+ (\hat{r}_{FEMALE} \times FEMALE) + (\hat{r}_{EDUC} \times EDUC) + (\hat{r}_{AGE} \times AGE) + (\hat{r}_{NHHD} \times NHHD),\end{aligned}$$

where $\hat{r}_0$ is the estimate of the constant, normalised on India in terms of countries. If we collapse this specification by dropping all individual characteristics and country dummies, we would simply be estimating the constant terms for r.

The estimates allow for the possibility of correlation between responses by the same subject, so the standard errors on estimates are corrected for the possibility that the eight responses are clustered for the same subject. The use of clustering to allow for 'panel effects' from unobserved individual effects is common in the statistical survey literature.[7] Our estimates also allow for the stratification of observations by village (and hence also country).

2.2. *Prospect Theory Specification*

There are two components to the PT specification, the utility function and the probability weighting function.

We use the same CRRA functional form as specified for EUT: $U(x) = (x^{1-\alpha})/(1-\alpha)$. We do not have any losses in the lotteries considered here, so we drop the part of the utility function in PT that is defined for losses. Our evaluation of EUT and PT is therefore restricted to the gain domain, which is the domain over which most of the initial tests of PT were conducted.

There are two variants of PT, depending on the manner in which the probability weighting function is combined with utilities. The original version proposed by Kahneman and Tversky (1979) posits some weighting function which is separable in

[7] Clustering commonly arises in national field surveys from the fact that physically proximate households are often sampled to save time and money but it can also arise from more homely sampling procedures. For example, Williams (2000; p. 645) notes that it could arise from dental studies that 'collect data on each tooth surface for each of several teeth from a set of patients' or 'repeated measurements or recurrent events observed on the same person.' The procedures for allowing for clustering allow heteroscedasticity between and within clusters, as well as autocorrelation within clusters. They are closely related to the 'generalised estimating equations' approach to panel estimation in epidemiology (Liang and Zeger, 1986) and generalise the 'robust standard errors' approach popular in econometrics (Rogers, 1993). Wooldridge (2003) reviews some issues in the use of clustering for panel effects, in particular noting that significant inferential problems may arise with small numbers of panels.

© The Author(s). Journal compilation © Royal Economic Society 2009

outcomes, and has been usefully termed Separable Prospect Theory (PT) by Camerer and Ho (1994; p. 185). The alternative version, proposed by Tversky and Kahneman (1992), posits a weighting function defined over cumulative probability distributions. In either case, the weighting function proposed by Tversky and Kahneman (1992) has been widely used. It is assumed to have well-behaved endpoints such that $w(0) = 0$ and $w(1) = 1$ and to imply weights $w(p) = p^{\gamma}/[p^{\gamma} + (1 - p)^{\gamma}]^{1/\gamma}$ for $0 < p < 1$. The normal assumption, backed by a substantial amount of evidence reviewed by Gonzalez and Wu (1999), is that $0 < \gamma < 1$. This gives the weighting function an 'inverse S-shape', characterised by a concave section signifying the overweighting of small probabilities up to a crossover-point where $w(p) = p$, beyond which there is then a convex section signifying underweighting. If $\gamma > 1$ the function takes the less conventional 'S-shape', with convexity for smaller probabilities and concavity for larger probabilities.

Assuming that PT is the true model, prospective utility (PU) is defined in much the same manner as when EUT is assumed to be the true model. The PT utility function is used instead of the EUT utility function and $w(p)$ is used instead of p but the steps are otherwise essentially identical (Harrison and Rutström, 2008; §3.1). The difference in prospective utilities is defined similarly as $\nabla PU = PU_R - PU_L$. Thus the likelihood, conditional on the PT model being true, depends on the estimates of α and γ given the above specification and the observed choices. The conditional log-likelihood is

$$\ln L^{PT}(\alpha, \gamma; y, \mathbf{X}) = \sum_i l_i^{PT} = \sum_i [(\ln \Psi(\nabla PU) \times \mathbf{I}(y_i = 1)) + (\ln(1 - \Psi(\nabla PU)) \times \mathbf{I}(y_i = 0))].$$

The parameters α and γ can again be estimated as linear functions of the vector $\mathbf{X}$.

The use of probability weighting introduces another way in which individuals might be averse to risk, quite apart from the implied aversion to risk from having a concave utility function. The idea that one could use non-linear transformations of the probabilities in a lottery when weighting outcomes, instead of non-linear transformations of the outcome into utility, was most sharply presented by Yaari (1987). To illustrate the point clearly, he assumed a linear utility function, in effect ruling out any risk aversion or risk seeking from the shape of the utility function *per se*. Instead, concave (convex) probability weighting functions would imply risk seeking (risk aversion). In general, in PT one can have aversion to risk from either or both of probability weighting and utility curvature. We follow EUT convention and refer to 'risk aversion' or 'risk attitude' solely as a property of the curvature of the utility function, even when we refer to PT estimates, but this semantic point should be kept in mind and we return to it when it affects the interpretation of results.

2.3. *A Mixture Model Specification*

If we let π^{EUT} denote the probability that the EUT model is correct, and $\pi^{PT} = (1 - \pi^{EUT})$ denote the probability that the PT model is correct, the grand likelihood can be written as the probability weighted average of the conditional likelihoods. Thus the likelihood for the overall model estimated is defined by

$$\ln L(r, \alpha, \gamma, \pi^{EUT}; y, \mathbf{X}) = \sum_i \ln[(\pi^{EUT} \times l_i^{EUT}) + (\pi^{PT} \times l_i^{PT})].$$

© The Author(s). Journal compilation © Royal Economic Society 2009

This log-likelihood can be maximised to find estimates of the parameters. Just as we allowed the parameters for EUT and PT to be estimated as linear functions of the observables **X**, we could do so in this case. However, the sample is not sufficiently large to allow robust estimation of the mixture model with the full set of covariates, so we restrict analysis to including the country dummies for π^{EUT}.

3. Results

3.1. *Estimates of the EUT Specification*

Table 2 collates the estimates from our data assuming that EUT is the sole theory explaining behaviour. Panel (*a*) presents estimates assuming no covariates and panel (*b*) extends this by including covariates. Figure 1 displays the predicted distribution of risk attitudes, using the estimated model that includes covariates. This distribution reflects the predicted values of the CRRA coefficient *r*, where the prediction depends on the characteristics of the individual, the location of the experiments, and the order in which the tasks were presented. These results point to moderate risk aversion over these stakes, with virtually no evidence of any risk-loving behaviour in the sample as a whole. In Panel (*a*) the coefficient of CRRA is estimated to be 0.536, remarkably close to estimates obtained with comparable experiments and statistical methods in developed countries.[8]

From Panel (*b*) we observe that, compared to the model with no covariates, estimated risk aversion is slightly higher on average in India (0.841), which is the implicit country captured by the constant term. It is 0.050 higher in Ethiopia but this effect only has a p-value of 0.502; it is 0.169 higher in Uganda and this effect has a p-value of 0.015. The order of experimental tasks had a mild effect on elicited risk attitudes. Women appear to be slightly less risk averse than men, although the quantitative effect is small (−0.085) and barely significant (p-value = 0.068). There is a statistically significant effect from age but it is quantitatively small: every 10 years of age is associated with a decline in risk aversion by 0.06.

Table 2
Maximum Likelihood Estimates of EUT Model of Choices

Coefficient	Variable	Estimate	Standard Error	p-value	95% Confidence Intervals	
(*a*) *No Covariates*						
r	Constant	0.536	0.024	0.000	0.488	0.583
(*b*) *Including Covariates*						
r	Constant	0.841	0.091	0.000	0.662	1.021
	Ethiopia	0.050	0.074	0.502	−0.095	0.195
	Uganda	0.169	0.070	0.015	0.032	0.306
	Order of tasks	−0.063	0.059	0.283	−0.178	0.052
	Age in years	−0.006	0.002	0.002	−0.010	−0.002
	Female	−0.085	0.046	0.068	−0.176	0.006
	Some secondary education	0.056	0.059	0.345	−0.060	0.172
	Number in household	−0.013	0.010	0.178	−0.033	0.006

[8] For example, Holt and Laury (2002) and Harrison, Johnson, McInnes and Rutström (2005) for college students in the US and Harrison, Lau and Rutström (2007) and Harrison, Lau, Rutström and Sullivan (2005) for the adult population in Denmark.

© The Author(s). Journal compilation © Royal Economic Society 2009

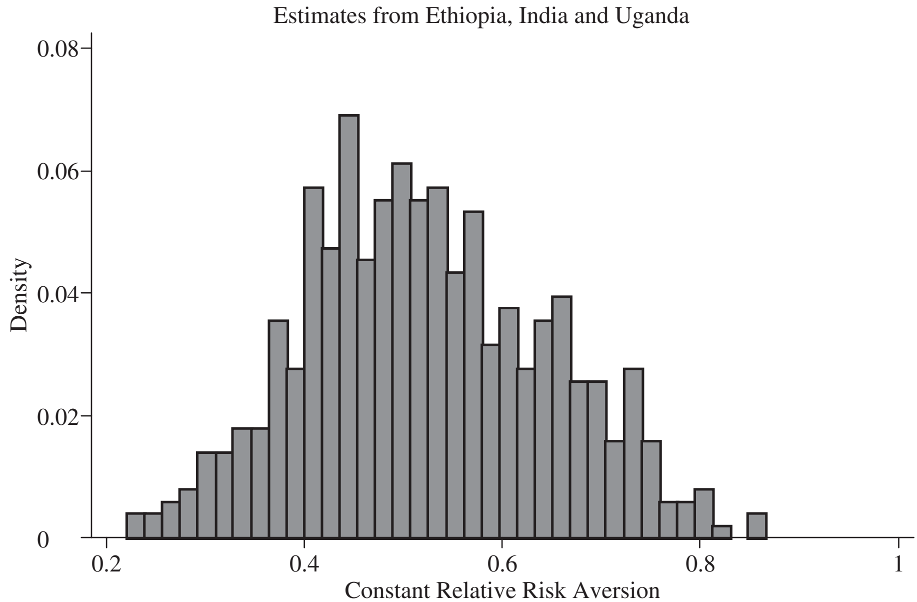

Fig. 1. *Risk Attitudes Assuming EUT*

3.2. *Estimates of the PT Specification*

Table 3 collates estimates of the data assuming that PT is the sole model explaining the data, and that we use the probability weighting function proposed by Tversky and Kahneman (1992). Again, Panel (*a*) shows estimates that apply one model to all subjects and Panel (*b*) includes covariates for each parameter. From Panel (*a*) we see that the estimates of risk attitudes are considerably lower than under EUT ($\alpha = 0.464 < 0.536 = r$), although still consistent with risk aversion since $\alpha > 0$. The explanation, of course, is that probability weighting is allowed and that substitutes for some of the concavity of the utility function when explaining the data. The estimate of γ from Panel (*a*) of Table 3 is 1.384. This implies an S-shaped probability weighting function which would entail underweighting of low probabilities and overweighting of large probabilities. This result contrasts starkly with the empirical claims from data collected in experimental laboratories in developed countries (Gonzalez and Wu, 1999). The left panel in Figure 2 illustrates the function implied by this estimate: it clearly has underweighting for probabilities below 0.6 but the extent of overweighting for higher probabilities is not great.[9] We return to consider more flexible functions below.

Including covariates in the PT specification leads to qualitative conclusions about risk attitudes that are similar to those obtained under EUT. Subjects in Ethiopia are estimated to be slightly more risk averse than those in India (+0.033) but the effect is

[9] The bottom axis of each panel in Figure 2 shows the probability that was presented to the subject in a task and the vertical axis shows the estimated weighted probability that the subject used. Overweighting means that the subject has a $w(p)$ estimate that is greater than the p it corresponds to; underweighting is the reverse situation in which $w(p) < p$.

© The Author(s). Journal compilation © Royal Economic Society 2009

Table 3

Maximum Likelihood Estimates of PT Model of Choices With Tversky-Kahneman Probability Weighting Function

Coefficient	Variable	Estimate	Standard Error	p-value	95% Confidence Intervals	
(*a*) *No Covariates*						
α	Constant	0.464	0.036	0.000	0.393	0.535
γ	Constant	1.384	0.070	0.000	1.246	1.522
(*b*) *Including Covariates*						
α	Constant	0.896	0.128	0.000	0.645	1.147
	Ethiopia	0.033	0.126	0.792	−0.215	0.281
	Uganda	0.195	0.097	0.045	0.004	0.385
	Order of tasks	−0.056	0.098	0.566	−0.248	0.136
	Age in years	−0.007	0.003	0.031	−0.014	−0.001
	Female	−0.104	0.071	0.144	−0.245	0.036
	Some secondary education	0.061	0.081	0.448	−0.097	0.219
	Number in household	−0.027	0.016	0.094	−0.060	0.005
γ	Constant	0.690	0.375	0.066	−0.046	1.427
	Ethiopia	0.045	0.396	0.909	−0.734	0.824
	Uganda	0.208	0.169	0.219	−0.124	0.540
	Order of tasks	0.211	0.165	0.201	−0.113	0.535
	Age in years	0.004	0.012	0.742	−0.019	0.027
	Female	0.002	0.187	0.993	−0.366	0.370
	Some secondary education	0.076	0.159	0.634	−0.236	0.387
	Number in household	0.060	0.036	0.092	−0.010	0.131

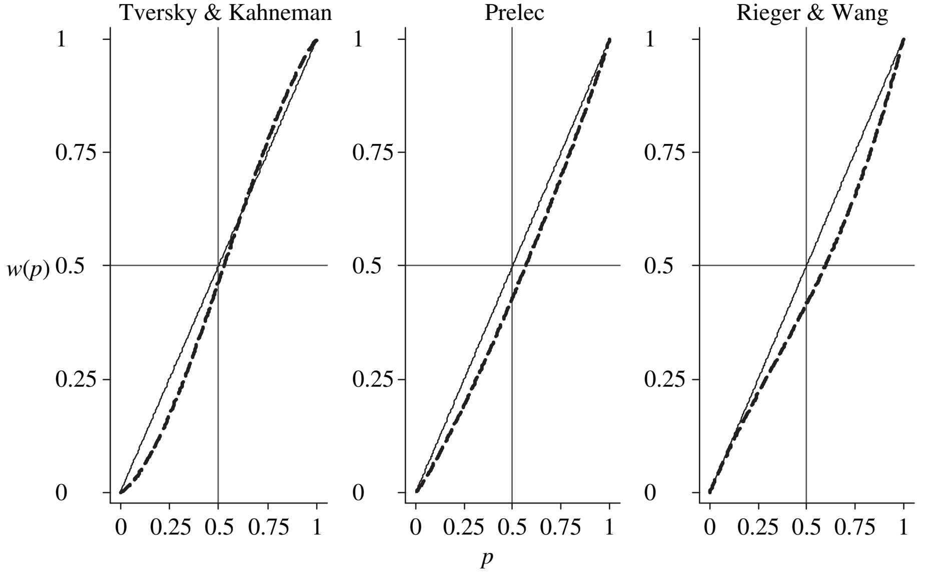

Fig. 2. *Alternative Probability Weighting Functions*

not statistically significant (p-value = 0.792). However, those in Uganda are estimated to be much more risk averse on average (+0.195), and the effect is significant (p-value = 0.045). Women are again slightly less risk averse than men and the effect is

© The Author(s). Journal compilation © Royal Economic Society 2009

barely significant (p-value $= 0.144$). The effect of age is roughly the same as for EUT. There does not appear to be major differences in the extent of probability weighting across countries. The size of the household does affect the average extent of probability weighting.

There are some limitations of the conventional Tversky and Kahneman (1992) probability weighting function. It does not allow independent specification of location and curvature; it has a fixed point, where $p = w(p)$ at $p = 1/e = 0.37$ for $\gamma < 1$ and at $p = 1 - 0.37 = 0.63$ for $\gamma > 1$; and it is not even increasing in p for small values of γ. Prelec (1998) offers a two-parameter probability weighting function that exhibits more flexibility than the Tversky and Kahneman (1992) function. The Prelec (1998) function is $w(p) = \exp[-\eta(-\ln p^{\phi})]$, which is defined for $0 < p < 1$, $\eta > 0$ and $0 < \phi < 1$. Rieger and Wang (2006; Proposition 2) offer a two-parameter polynomial of third degree which is defined for $0 \leq p \leq 1$, unlike the Prelec (1998) function: $w(p) = p + [(3 - 3b)/(a^2 - a + 1)][p^3 - (a + 1)p^2 + ap]$, where $0 < a < 1$ and $0 < b < 1$. The parameter restrictions on a and b ensure that the function is concave for lower values of p and then convex for larger values of p. Values of b larger than 1 would allow convex and then concave shapes, which we want to allow *a priori* given the findings of Humphrey and Verschoor (2004*a,b*).[10]

Table 4 reports estimates of the PT model assuming these two alternative functional forms; Figures 2 and 3 display the effects on the shape of the probability weighting function and elicited risk attitudes. Both of the alternatives confirm the presence of significant underweighting of probabilities over a wide range of probabilities. In fact, both of the two-parameter probability weighting functions are *weakly* well-behaved with respect to the conventional empirical wisdom that there should be a concave and then convex ('inverse-S') shape. These shapes, in fact, are quite close to the original form sketched by Kahneman and Tversky (1979; p. 283), which exhibited considerable under-weighting of probabilities for virtually the whole range of p.

We observe from Table 4 and Figure 3 that the implied risk attitudes are mildly sensitive to the use of the two flexible probability weighting functions. The Prelec (1998) function leads to estimates that are mid-way between those obtained with the Tversky and Kahneman (1992) function and EUT, and the Rieger and Wang (2006) function leads to estimates that are closer to EUT. These differences derive from seemingly small differences in the probability weighting functions shown in Figure 2: the Prelec (1998) function exhibits the greatest under-weighting for lower probabilities and the Rieger and Wang (2006) function exhibits the greatest under-weighting for all probabilities.

Since it matters for inferences about risk attitudes, how should one select from these two flexible functional forms? The Prelec (1998) function is implied by a series of properties that it is claimed that the function has to satisfy, many of which have been inferred from previous experimental tasks rather than from theoretical considerations. There is nothing wrong with this procedure, apart from the fact that it rests as a logical matter on those prior empirical inferences being valid. The Rieger and Wang (2006) function, on the other hand, is derived as the simplest polynomial to satisfy some theoretically attractive properties, most notably that it be strictly increasing and

[10] In this case b must not exceed $1 + 1/3[(a^2 - a + 1)/(1/2 + |a - 1/2|)]$ or the function becomes non-increasing (Marc Oliver Rieger; personal communication).

© The Author(s). Journal compilation © Royal Economic Society 2009

Table 4

Maximum Likelihood Estimates of PT Model of Choices With Different Probability Weighting Function

Coefficient	Estimate	Standard Error	p-value	95% Confidence Intervals	
(*a*) *Tversky & Kahneman Probability Weighting Function:* $w(p) = p^{\gamma}/[p^{\gamma} + (1-p)^{\gamma}]^{1/\gamma}$					
α	0.464	0.036	0.000	0.393	0.535
γ	1.384	0.070	0.000	1.246	1.522
(*b*) *Prelec Probability Weighting Function:* $w(p) = \exp[-\eta(-\ln p^{\phi})]$					
α	0.504	0.033	0.000	0.439	0.569
η	1.202	0.053	0.000	1.097	1.307
ϕ	0.963	0.076	0.000	0.814	1.113
(*c*) *Rieger & Wang Probability Weighting Function:* $w(p) = p + [(3-3b)/(a^2 - a + 1)][p^3 - (a+1)p^2 + ap]$					
α	0.546	0.025	0.000	0.496	0.596
a	0.000	†	†	†	†
b	0.775	0.048	0.000	0.680	0.870
(*d*) *Estimates of Risk Aversion Parameter* α *Using the Rieger & Wang Probability Weighting Function*					
Constant	0.823	0.102	0.000	0.622	1.024
Ethiopia	0.055	0.081	0.500	−0.104	0.213
Uganda	0.198	0.093	0.034	0.015	0.381
Order of tasks	−0.057	0.065	0.382	−0.186	0.071
Age in years	−0.006	0.002	0.012	−0.011	−0.001
Female	−0.084	0.050	0.098	−0.183	0.015
Some secondary education	0.053	0.062	0.394	−0.069	0.175
Number in household	−0.015	0.013	0.265	−0.040	0.011

†The point estimate for a is 1.60e−28. It is not possible to calculate estimates of the standard error because of the lack of numerical precision at such extreme values. Parameter a is estimated by estimating a non-linear transform $\kappa \in (-\infty, +\infty)$, where $a = 1/[1 + \exp(\kappa)]$. Then the point estimates and standard errors of a are recovered from the estimates for κ using the 'delta method', which requires that derivatives be calculated in the neighbourhood of the point estimate. For certain extreme values of these point estimates, these numerical derivatives become unstable and the estimated standard error unreliable.

continuously differentiable on $p \in [0,1]$. Thus it does not depend, for it is *a priori* validity, on the validity of prior empirical tests, and for our purposes is preferable.[11]

Figure 4 compares directly the distribution of elicited risk attitudes in our sample under EUT or PT. These distributions are based on predictions from estimated models that include the full set of covariates for each parameter. Thus we use the predictions from the estimates shown in Panel (*b*) of Table 2 and Panel (*d*) in Table 4. Although there is a slight increase in estimates of risk aversion under PT, the results are remarkably similar, and the slight non-linearity of the probability weighting function would not change that conclusion.

3.3. *Estimates of the Mixture Model*

Finally, we extend the comparison of the two models to consider the mixture model that allows both to play a role in explaining observed behaviour. We employ the EUT model and the PT model with the Rieger and Wang (2006) probability weighting function. Maximum likelihood estimates are reported in Table 5.

[11] Underlying this perspective is some agnosticism with respect to assertions that previous tests of EUT provide clear conclusions that are independent of complications. From different perspectives, Harrison (1994), Humphrey (2000) and Harrison and Rutström (forthcoming) illustrate our concerns.

© The Author(s). Journal compilation © Royal Economic Society 2009

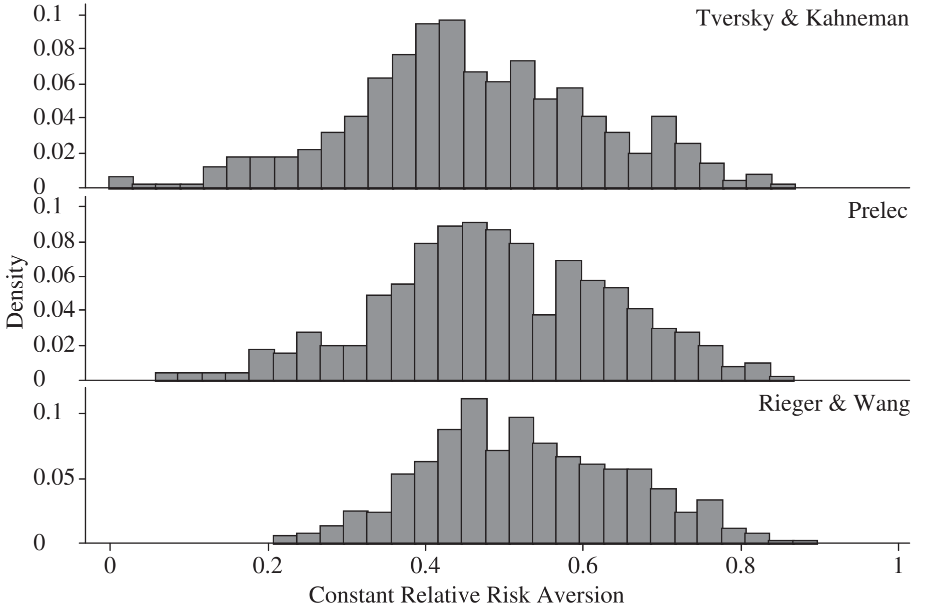

Fig. 3. *Alternative Risk Attitudes Assuming Prospect Theory*

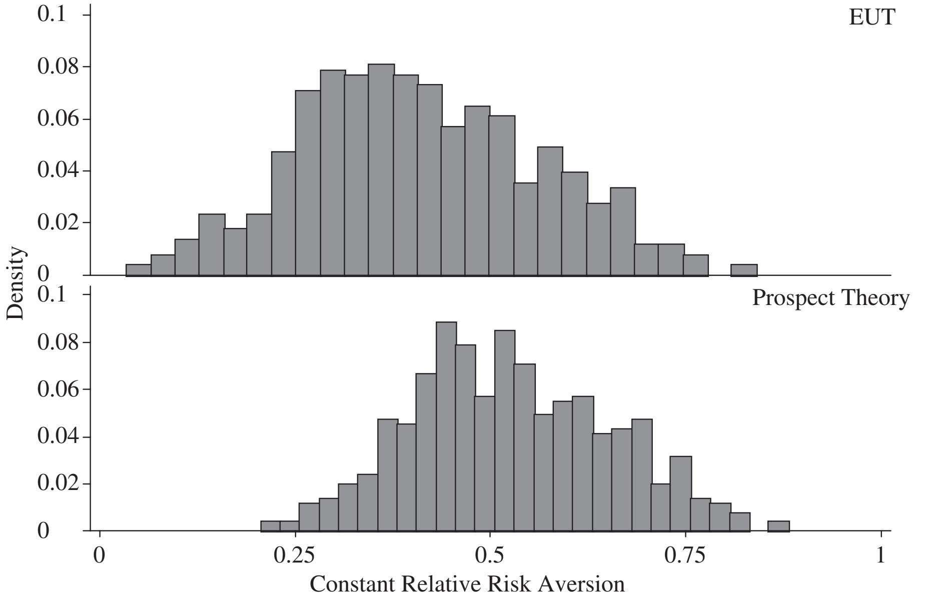

Fig. 4. *Risk Attitudes Under EUT or PT*

The first result is that the estimated probability for the EUT model is 0.461, and that this estimate is significantly different from 0 or 1 (p-value < 0.001). A test of the null hypothesis that $\pi^{EUT} = 1/2$ has a p-value of 0.44 and the upper and lower bounds of the 95% confidence interval for π^{EUT} are 0.36 and 0.56. Thus we *might* be inclined to

© The Author(s). Journal compilation © Royal Economic Society 2009

Table 5

Maximum Likelihood Estimates of Mixture Model

Coefficient	Estimate	Standard Error	p-value	95% Confidence Intervals	
α	−0.195	0.061	0.001	−0.315	−0.076
a	9.11e−08	†	†	†	†
b	4.09e−07	†	†	†	†
r	0.796	0.035	0.000	0.727	0.866
π^{EUT}	0.461	0.050	0.000	0.363	0.559

† It is not possible to calculate estimates of the standard error of a and b because of the lack of numerical precision at such extreme values. Parameter a is estimated by estimating a non-linear transform $\kappa \in (-\infty,+\infty)$, where $a = 1/[1 + \exp(\kappa)]$; a similar transform is used for parameter b. Then the point estimates and standard errors of a are recovered from the estimates for κ using the 'delta method', which requires that derivatives be calculated in the neighbourhood of the point estimate. For certain extreme values of these point estimates, these numerical derivatives become unstable and the estimated standard error unreliable.

conclude that the weight of the evidence supports PT over EUT by a (quantum) nose, but that would be an invalid inference for reasons explained earlier. Instead, we conclude that the data is consistent with each model playing a roughly equal role as a data generating process.

We believe that this finding is of more general significance. For example, Humphrey (2000; p.260) draws the following conclusion from some common consequence tests with subjects from developed countries that resonates well with our findings:

> The data are not explained by any of the generalised expected utility models which were developed to explain observed violations of expected utility theory in decision-problems of exactly the type used in this experiment. More worrying, perhaps, is that minor changes in problem representation seemingly impart large changes in choices. It is not surprising, therefore, that any single model is descriptively inadequate. Starmer (1992; p. 829) suggests that individual choice behaviour is 'more subtle and complex' than decision theorists have generally conveyed in their models. If so, this may render the induction of theories from sub-sets of experimental evidence problematic. [...]
>
> This conclusion depends upon the perceived role of theory. If a *single* theory should explain as much (as parsimoniously) as possible, the volumes of diverse observed influences on decision-making behaviour seemingly condemn this task to inevitable failure. If, however, risky choice is recognised as being too complex to be captured by any single theory and that the role of a single theory is to capture a *facet* of behaviour in a *specific* context, then it may be necessary to accept that slightly different contexts will invoke additional facets of behaviour and overall explanations of data will require more than one model. Although the behaviour observed in this experiment *might*, with sufficient ingenuity, be explained by a single model involving a complex probability weighting function, experience suggests that any such function will be limited in its application to other types of decision problem.

Our results are also consistent with more elaborate mixture models applied to larger databases of choice under uncertainty by Harrison and Rutström (forthcoming).

© The Author(s). Journal compilation © Royal Economic Society 2009

The second result from Table 5 is that the estimated risk attitudes and shape of the estimated probability weighting function change significantly under PT. In fact, the sample exhibits significant *risk-loving* behaviour (a convex utility function) to the extent that it follows the PT data generating process ($\alpha = -0.195$) but remains risk averse to the extent that it follows the EUT data generating process ($r = +0.796$). In interpreting this result it is important to emphasise that, as discussed above, this result refers only to the curvature of the utility function. Moreover, since the influence of the convex probability weighting function in Figure 5 will be to partially offset this risk seeking, the net effect on behaviour will be to make it more similar to that of the EUT decision-makers than the different curvatures of the utility function would imply.

This second point can be illustrated by using the parameters of the PT model to infer a risk premium in the traditional sense and, hence, infer a certainty-equivalent (Levy and Levy, 2002). That certainty-equivalent can then be compared directly to the standard certainty-equivalent under EUT. Using the estimates from Panel (*a*) of Tables 2 and 3, one can show that these certainty-equivalents only differ by about 6% for most of the lotteries in our experiment. An Appendix shows the results of a simulation of this method using the estimates from Panel (*a*) and Panel (*b*) of Tables 2 and 3.

If, as these simulation results suggest, the revealed behaviour of EUT decision-makers is broadly similar to that of PT decision-makers, then the question arises as to how one might interpret the results of the mixture model analysis in a behavioural sense. We suggest that a natural interpretation is that the mixture model estimates are identifying the different decision-making processes that are used by different individuals to arrive at these, broadly similar, choices.

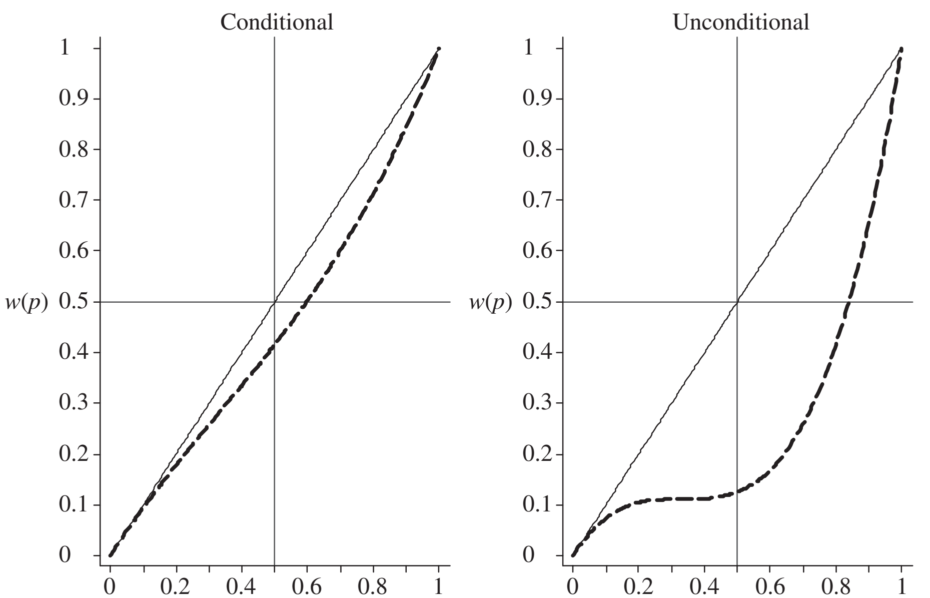

Fig. 5. *Effect of Mixture Model on Probability Weighting*

© The Author(s). Journal compilation © Royal Economic Society 2009

Although the mixture model applied to our data does not allow us to identify the precise nature of these decision-making processes, there are several plausible possibilities suggested in the literature. Brandstätter *et al.* (2002) argue that emotions are important in determining the shape of the probability weighting function. The underweighting of probabilities that we observe, for example, is consistent with pessimism regarding chance events. The mixture model could then be interpreted, for any given and broadly similar behaviour, as sorting out the pessimistic decision-makers from those who are neither pessimistic nor optimistic. The heterogeneity in PT decision-makers' utility functions might then be subsequently interpreted as reflecting 'residual' and different influences on risk-attitude.

For example, against the background of a convex probability weighting function, PT decision-maker *A* may be influenced by a factor *x* which is reflected in a convex utility function and still behave in the same manner as PT decision-maker *B* who is influenced by a different factor *y* which is reflected in a concave utility function. The only requirement is that PT decision-maker *A*'s convex utility function does not offset the convexity of the probability weighting function to the extent which would causes their behaviour to deviate from PT decision-maker *B*'s.

This example also suggests that it is perhaps unsurprising that PT decision-makers are more heterogeneous in their utility functions than EUT decision-makers. The only difference between our PT specification and EUT specification is that the former involves an additional process which is captured by the probability weighting function. For given behaviour, this allows heterogeneity in utility for PT decision-makers which is simply not possible for EUT decision-makers. Indeed, PT allows for four possible effects on risk attitudes from these two distinct processes: agents can be more or less risk averse because of the shape of their utility function, *ceteris paribus* their decision weights, and they can be more or less risk averse because of the shape of their probability weighting function, *ceteris paribus* their utility function.[12] Thus we have some basis for believing that the heterogeneity in risk attitudes we see with our PT specification should be a general outcome.

The effect of allowing for the mixture model estimates on the probability weighting function are illustrated in Figure 5. The panel on the left shows the estimated function when PT was assumed to be the sole data generating process and 'had to' explain 100% of these data.[13] The panel on the right of Figure 5 shows the estimated function when PT only has to account for 54% of these data and EUT is allowed to explain the other 46% of these data. It exhibits the same qualitative shape as the function estimated conditional on PT being the sole data generating process but with a marked increase in the underweighting of probabilities.

These results also force one to pay attention to the choice of parametric models for utility and probability weighting. The Reiger and Wang (2006) function is actually 'well-behaved' with the parameter values in Table 5: even though it has a proximately flat region for probabilities between 0.2 and 0.4, it is strictly increasing, weakly concave for the lowest probabilities and then sharply convex for most of the probabilities used in the lotteries.

[12] If one also allowed for loss aversion, in a design that allowed for losses, there would be a third process that could additionally allow for two additional effects on risk attitudes.

[13] The panel on the left of Figure 5 is the same function shown on the far right panel of Figure 2.

© The Author(s). Journal compilation © Royal Economic Society 2009

We also extended the mixture model to include binary dummies for Ethiopia and Uganda for the π^{EUT} parameter. The results indicate that there is least support for the EUT model on average in India (0.35) than in Ethiopia (0.57) or Uganda (0.51). We can reject the hypothesis that the EUT and PT models have equal explanatory power in India (p-value = 0.014) but not in Ethiopia (p-value = 0.38) or Uganda (p-value = 0.85). However, even in the case of India, the 95% confidence intervals for the support of the EUT model are between 0.23 and 0.47.

The changes in results under PT are striking as we move from the original specification to the mixture model. Consider the underweighting of probabilities. Underweighting means that when subjects are told that some outcome has a 50% chance of occurring that they behave as if it has much less chance of occurring. This appears to be true for all of the probabilities in our lotteries, which range from 1/4 to 3/4. One possible explanation for this observation is that our mixture model estimates reflect the *pessimism* of PT subjects. Our subjects might behave pessimistically because of the general economic conditions prevailing at the time of the experiments. The regions we visited in India and Ethiopia were experiencing droughts. If this served to engender a general pessimism about uncertain events, this might account for our results.

There is some support for this interpretation from psychological studies of the same phenomenon. Hertwig *et al.* (2004) review evidence from a range of psychology experiments, some with real rewards and without deception (Barron and Erev, 2003), that provide striking evidence of underweighting at low probabilities. They argue that this is characteristic of tasks that involve 'decisions from experience', where the probabilities in question derive from the subject's previous experience with comparable events.[14] They also argue that evidence of overweighting at low probabilities derives from tasks that involve 'decision from description', where the task description itself provides the probabilities. In turn, they argue that the underweighting behaviour might derive from reliance on relatively small samples of information and the overweighting of recently sampled information. Our experiments do not allow us to discriminate between these lines of argument but the significant underweighting we observe does suggest that our subjects viewed the probabilities in the task description in terms of their experience, perhaps from naturally-occurring risk in their environment.

When viewed in the light of the findings of existing studies, providing an interpretation of our evidence of the underweighting of probabilities brings into focus the complex issue of the influence of experience on decision-making. List (2004) argues that evidence from field experiments shows that decisions taken by experienced decision-makers can be organised by standard neoclassical economic theory, specifically, Hicksian consumer preferences. On the other hand, he argues that organisation of the data he observes from inexperienced decision-makers requires a reference-dependent utility function of the type employed in prospect theory. List (2004) studies decisions over whether to trade mugs for chocolate bars; if this evidence was applied to the risky lottery choices which we study, then it would imply that our prospect theory decision-makers, for whom we observe the underweighting of probabilities, may simply be less experienced in risky decision-making than the EU decision-makers. This implication

[14] This experience can also come from simply providing feedback to the subject about the mechanism used to make random draws. We did not provide such feedback in our experiments.

© The Author(s). Journal compilation © Royal Economic Society 2009

would be at odds with the previous interpretation suggested by Hertwig *et al.* (2004) and is not something we can resolve satisfactorily with the present data. It does point to the value of an important extension of our design to build in natural controls for experience with the type of decision being studied, as stressed in the general literature on field experiments (Harrison and List, 2004) as well as specific field experiments examining risk attitudes (Bohm and Lind, 1993; Harrison, List and Towe, 2007).

Another possible explanation for underweighting is that it reflects subjects not believing that the random process was actually fair, despite the fact that we used no deception whatsoever, used transparent physical randomisation devices and saw no evidence that our subjects were concerned with being cheated. However, if the subjects believe that the experimenter had a way of making the outcome actually go against them, then one might expect to see behaviour of this kind. Such concerns are always a part of any experiment, of course, and are the reason that many experimental economists use physical randomising devices rather than rely on computers whenever possible. But it is distinctly possible that cultural beliefs about certain physical randomising devices, and experiences with being 'cheated' in such interactions, are different in developing countries.[15]

Now consider the *qualitative* change in risk attitudes under the PT model when one moves from assuming it to be the only data-generating process to being one of two possible data-generating process. One might ask if this result is an artefact of the use of a mixture model. Intuitively, if EUT can explain about 50% of the sample data and, if all of these subjects happen to be risk averse, one might ask whether the mixture model simply assigns the risk-loving subjects to the PT model since there is no alternative model for them to be assigned to. Thus what appears to be a change in risk attitudes under PT is, according to this view, just due to the risk lovers being 'residually' assigned to the PT model. Although difficult to state formally, this is a good question, which goes to the heart of the use of statistical models to simultaneously identify parameters *and* alternative models. This question is in effect a comment on the potential dangers of assuming that the EUT decision-maker and the PT decision-maker are each homogeneous: they can have different risk attitudes in the specification estimated in Table 5, but if you are an EUT decision-maker you have to have the same risk attitude as every other EUT decision-maker.

A complete response to this question would require that one include individual characteristics of the respondents in the parameters of the mixture model, in order to identify any heterogeneity within the subset of EUT or PT decision-makers. We do not have sufficient degrees of freedom to do this for the whole model but we make the following two observations.

[15] The outcomes in our experiment were substantial and, in terms of PPP, above those typically offered in experiments in developed countries. In experiments similar to ours but conducted in developed countries, large incentives might be considered an attractive design feature: they motivate careful consideration of decisions and may offset some 'induced' risk seeking stemming from subjects 'gambling with the house money'. In our experiments large incentives may have induced the underweighting of probabilities if subjects believed that the opportunity to gamble with such large sums of 'house money' was too good to be true. Such belief may have undermined confidence in the authenticity of the random devices employed to resolve risk. This possibility is likely to be mitigated in developed countries where experimental subjects, who are usually undergraduate students, are immersed in an environment where being paid non-trivial sums to complete simple experimental tasks for research purposes is likely to be viewed less suspiciously.

© The Author(s). Journal compilation © Royal Economic Society 2009

First, we can identify heterogeneity within the group of PT decision-makers with respect to risk aversion coefficient α. The average value for this coefficient is -0.16, consistent with the estimate from the homogeneous-PT model in Table 5. But these estimates show a significant variation in the risk attitudes within the subset of PT decision-makers. Those in India are the most risk-loving on average (-0.32), with those in Uganda being risk-neutral on average (0.03) and those in Ethiopia being in-between and risk-loving (-0.16).[16] The presence of some secondary education is associated with a significantly higher aversion to risk at the margin (+0.14, with a p-value of 0.045) and there is a dramatic effect of age. Every additional year lowers risk aversion by 0.033 and this a marginal effect that is statistically significant (p-value = 0.002). The effect of age can be seen in Figure 6, which stratifies the predicted risk aversion coefficient α under PT for each subject. Younger subjects tend to be risk averse under PT and older subjects tend to be risk loving under PT. Whether or not the same effect is observed under EUT, Figure 6 dramatically illustrates that there is considerable variation in risk attitudes within the subset of PT decision-makers.

By way of contrast, Figure 7 shows a comparable display of the association of age on risk attitudes within the subset of EUT decision-makers. Although there is a similarly declining marginal effect on risk attitudes (-0.005 per year of age), the effect is not statistically significant (p-value = 0.56). Thus we see that there is considerable sensitivity of the demographic pattern of risk attitudes to the type of choice theory that best explains behaviour. Thus reliable policy inferences about age and risk attitude should condition on the heterogeneity of the type of decision-making model being used as well as the observable characteristic age.[17]

Second, as discussed earlier, the estimates of certainty equivalents for the lotteries used in our experiments show that our subjects evaluated the lotteries broadly similarly in a quantitative sense, irrespective of whether they were assumed to be PT or EUT decision-makers. The fact that the PT certainty equivalents were all lower, albeit marginally so, than the EUT certainty equivalents implies that the latter group made riskier choices than the former group. This implication follows from the mixture model sorting subjects into two groups which correspond to the two latent processes we assume, and according to the heterogeneity or homogeneity of their utility functions. That is, given the significant underweighting manifest in $w(\cdot)$, the PT utility function necessarily exhibits greater heterogeneity if the theory as a whole is to explain the broadly similar evaluations of lotteries which EUT requires only a single utility function to organise. In this respect the insight of the mixture model approach is that the estimates of the PT risk aversion parameter α are substantially more heterogeneous than the quantitative evaluations of the lotteries might suggest.

Of course, there are many extensions of our approach possible before one can draw definitive conclusions. More data always helps but for statistical inferences

[16] These estimates are average effects, including variations in age and sex, for example, in each country.

[17] For example, one of the most important risky choices that our subjects make in practice is whether or not to take up the production of lucrative but input-intensive cash crops, such as tomatoes and cabbages. Agricultural extension workers told us that it was primarily young men who gamble all on such crops, for a season or two, often to finance their eventual migration to a town or city. Our findings are certainly inconsistent with the presumption that this behaviour is due to risk-loving tendencies among the young and may have broader implications for micro-finance schemes that consider financing the growing of lucrative, input-intensive crops and for the targeting of agricultural extension.

© The Author(s). Journal compilation © Royal Economic Society 2009

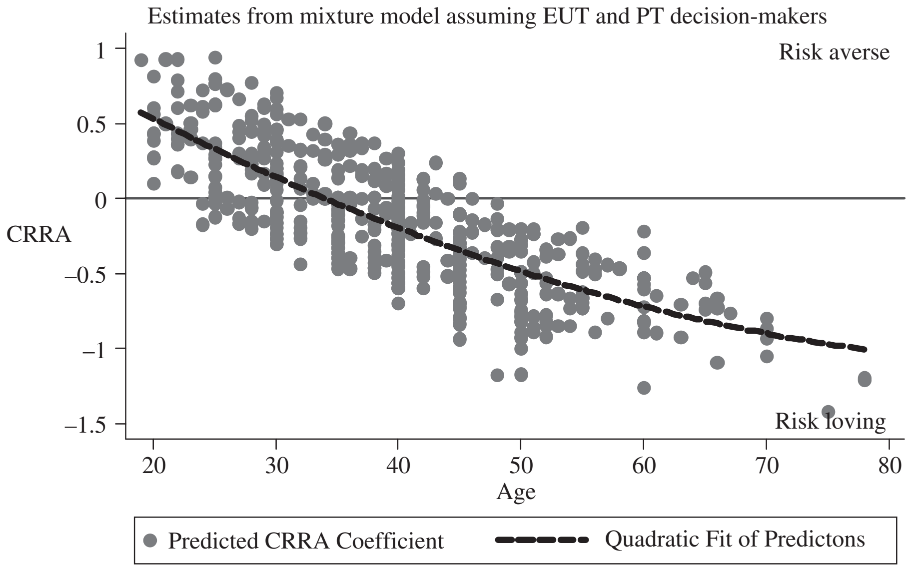

Fig. 6. *Effect of Age on Estimated Risk Aversion of PT Decision-Makers*

based on mixture models it is more than normally true since one remains 'agnostic' about which data-generating process dominates. This need for more data would only become more severe if one admitted more than two data-generating processes.[18] In addition, we would want to examine alternatives to EUT that have some theoretically attractive properties in comparison to the separable PT considered here. In this vein, it might be useful to examine some of the popular stochastic error specifications that have been proposed. We would also be particularly interested in extensions to explicitly consider outcomes frames as losses, to assess the effect of loss aversion on choices under uncertainty among the rural poor in developing countries.

Finally, we would encourage examination of risk-taking behaviour in the broader economics context appropriate for the country, village and individual: there is considerable evidence and theory to suggest that 'background risk' can influence risk-taking behaviour over 'foreground risk' (Harrison, List and Towe, 2007). Examples of this type of background risk include the effects of weather conditions, violence, corruption, reliability of government infrastructure, the efficacy of risk-sharing and risk-pooling arrangements at a social level and even norms for responses to individuals experiencing good luck or bad luck. We would expect there to be unusually wide variations in the extent of background risk and its composition in developing countries. One objective of our foreground risk task being standardised was to provide a controlled baseline against which these influences can be measured in future research.

[18] Intuitively, the need for data is relatively less severe if the alternative models have sharply different predictions and relatively more severe if the alternative models have similar predictions. We therefore doubt that one can easily discriminate between the competitors to EUT without significantly more controlled data, since they have many 'family similarities' (Starmer, 2000).

© The Author(s). Journal compilation © Royal Economic Society 2009

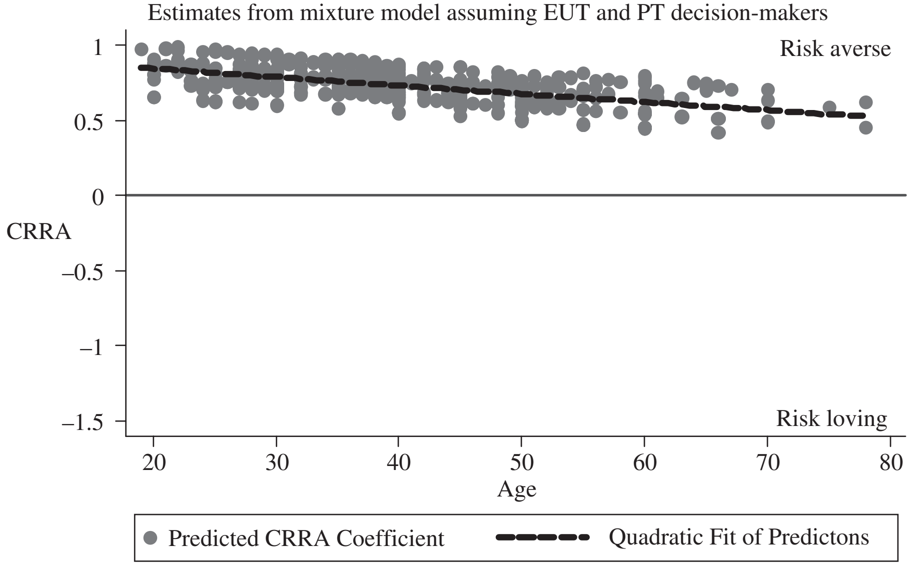

Fig. 7. *Effect of Age on Estimated Risk Aversion of EUT Decision-Makers*

4. Concluding Remarks

Our results show how important it is to be clear about the theoretical and statistical assumptions underlying inferences from observed data. For example, our results point to the dangers of drawing inferences about risk attitudes when one incorrectly assumes that behaviour is generated by only one data-generating process. When we do that and assume PT or EUT, we infer risk aversion and relatively little heterogeneity of risk attitudes. Moreover, our inferences about the degree of risk aversion do not appear to be affected by which of the two models we adopt. But when we allow some of the data to be explained by EUT and some to be explained by PT, we infer relatively homogeneous risk aversion for the subjects following EUT and considerable heterogeneity of risk attitudes for the subjects following PT. In fact, the heterogeneity of risk attitudes for the subjects following PT spans a significant fraction of risk-loving, risk-neutral and risk-averse individual behaviour. The average subject following PT exhibits risk-loving behaviour but the key finding here is heterogeneity of risk attitudes. The differences obtained when one allows multiple data-generating processes is a general point that is true for developed countries as well as developing countries. It may be more significant in developing countries where one might expect more noise in the data due to the relative unfamiliarity of the tasks and the adoption of a wider range of heuristics to make decisions.

Substantively, we conclude that there is roughly equal support for the two major models of choice under uncertainty considered here. It is not the case that EUT or PT wins but that the data is consistent with each playing some roughly equal role, as in Harrison and Rutström (forthcoming) for comparable laboratory tasks with university students. Thus, substituting PT for EUT would be tantamount to replacing one ‘half

© The Author(s). Journal compilation © Royal Economic Society 2009

wrong' assumption with another. This conclusion implies that policies should not be designed under the assumption that one *or* other theory explains all behaviour. Policy-makers are therefore faced with the substantial challenge of examining the sensitivity of the predicted impact of policy interventions to the replacement of single key assumptions with a probability distribution over competing assumptions, as suggested by estimates emerging from mixture model analyses based on experimental data. There is a clear a rationale for closer collaboration between experimentalists and policy-makers in developing countries than has hitherto been the case.

Appendix A: Measures of Risk Premia

Levy and Levy (2002) show that Pratt's measure of the risk premium is given by the ϱ which solves $\Sigma(p_i)u(x_i) = u[\Sigma w(p_i)x_i - \varrho]$. We can calculate this measure for the lotteries used in our experiment by implementing this formula with the CRRA utility function and the Tversky and Kahneman (1992) probability weighting function we assume. In this case the estimates of α and γ from Table 3 are used (columns *A* and *B*). We do the same on the assumption of EUT but in this case impose the restriction $\gamma = 1$ (column B) and use the estimate of r (column *A*) from Table 2. We can then add these calculated risk premia to the expected value of each lottery to provide a certainty equivalent value for that lottery, under the respective assumptions of EUT and PT (column *D*). Expressing the difference between these two certainty equivalents (column *E*) as a proportion of the EUT certainty equivalent provides a measure of the evaluation of each lottery by our PT decision makers in relation to our EUT decision makers (column *F*). We restrict this exercise to two outcome lotteries and report the results in the table below.

Pratt's Measure of Risk Premium

	A	*B*	*C*	*D*	*E*	*F*
Lottery (Table 1)	α or r	γ	ϱ	$p_ix_i+\varrho$	$(p_ix_i+\varrho)_{EUT}-(p_ix_i+\varrho)_{PT}$	$\frac{(p_ix_i+\varrho)_{EUT}-(p_ix_i+\varrho)_{PT}}{(p_ix_i+\varrho)_{EUT}}$
2A	0.464	1.384	33.29	95.79		
	0.536	1	49.91	112.41	16.62	14.8%
3A	0.464	1.384	38.23	225.73		
	0.536	1	53.29	240.79	15.07	6.3%
4A	0.464	1.384	123.70	398.70		
	0.536	1	151.53	426.53	27.82	6.5%
7A	0.464	1.384	66.57	191.57		
	0.536	1	99.79	224.78	33.21	14.8%
2B	0.464	1.384	22.53	72.54		
	0.536	1	27.55	77.55	5.01	6.5%
3B	0.464	1.384	7.83	182.83		
	0.536	1	9.13	184.13	1.3	0.7%
7B	0.464	1.384	22.95	135.45		
	0.536	1	31.84	144.34	8.89	6.2%

Technical Appendix

Additional supporting information may be found in the online version of this article and on the RES website:

Appendix B: Experimental Procedures

© The Author(s). Journal compilation © Royal Economic Society 2009

Please note: Wiley-Blackwell are not responsible for the contest or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Georgia State University and CEAR,
University of Osnabrück,
University of East Anglia

Submitted: 25 May 2006
Accepted: 24 November 2008

References

Barr, A. and Genicot, G. (2008). 'Risk sharing, commitment, and information: an experimental analysis', *Journal of the European Economic Association*, vol. 6(6), pp. 1151–85.

Barr, A. and Packard, T. (2002). 'Revealed preference and self insurance: can we learn from the self employed in Chile', Policy Research Working Paper No. 2754, World Bank, Washington DC.

Barr, A. and Packard, T. G. (2005). 'Seeking solutions to vulnerability in old age: preferences, constraints, and alternatives for coverage under Peru's pension system', Working Paper No. 2005-05, Centre for the Study of African Economies, Department of Economics, University of Oxford.

Barron, G. and Erev, I. (2003). 'Small feedback-based decisions and their limited correspondence to description-based decisions', *Journal of Behavioural Decision Making*, vol. 16, pp. 215–33.

Binswanger, H. P. (1980). 'Attitudes toward risk: experimental measurement in rural India', *American Journal of Agricultural Economics*, vol. 62 (August) pp. 395–407.

Binswanger, H. P. (1981). 'Attitudes toward risk: theoretical implications of an experiment in rural India', Economic Journal, vol. 91 (December) pp. 867–90.

Binswanger, H. P. (1982). 'Empirical estimation and use of risk preferences: discussion', *American Journal of Agricultural Economics*, vol. 64 (May) pp. 391–3.

Bohm, P. and Lind, H. (1993). 'Preference reversal, real-world lotteries, and lottery-interested subjects', *Journal of Economic Behavior and Organization*, vol. 22, pp. 327–48.

Botelho, A., Harrison, G. W., Pinto, L. M. C., Rutström, E. E. and Veiga, P. (2005). 'Discounting in developing countries: a pilot experiment in Timor-Leste', Working Paper No. 05-31, Department of Economics, College of Business Administration, University of Central Florida.

Brandstätter, E., Kühberger, A. and Schneider, F. (2002). 'A cognitive-emotional account of the shape of the probability weighting function', *Journal of Behavioral Decision Making*, vol. 15, pp. 19–100.

Camerer, C. F. (1989). 'An experimental test of several generalized utility theories', *Journal of Risk and Uncertainty*, vol. 2, pp. 61–104.

Camerer, C. F. (1995). 'Individual Decision Making', in (J.H. Kagel and A.E. Roth, eds.), *The Handbook of Experimental Economics*, pp. 587–704, Princeton: Princeton University Press.

Camerer, C. (1998). 'Bounded rationality in individual decision making', *Experimental Economics*, vol. 1, pp. 163–83.

Camerer, C. and Ho T. (1994). 'Violations of the betweenness axiom and nonlinearity in probabilities', *Journal of Risk and Uncertainty*, vol. 8, pp. 167–96.

Cardenas, J. C. and Carpenter, J. P. (2008). 'Behavioural development economics: lessons from field labs in the developing world', *Journal of Development Studies*, vol. 44(3), pp. 311–38.

Collier, P. and Gunning, J. W. (1999). 'Explaining African economic performance', *Journal of Economic Literature*, vol. 37, pp. 64–111.

Cox, J. C. and Sadiraj, V. (2006). 'Small- and large-stakes risk aversion: implications of concavity calibration for decision theory', *Games & Economic Behavior*, vol. 56(1), pp. 45–60.

Deaton, A. (1997). *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*, Baltimore: Johns Hopkins University Press and the World Bank.

Duflo, E. (2006). 'Field experiments in development economics', in (R. Blundell, W. Newey, and T. Persson, eds.), *Advances in Economics and Econometrics*, vol. 2, pp. 322–48, New York: Cambridge University Press.

Fafchamps, M. (2004). *Rural Poverty, Risk and Development*, Northampton, MA: Edward Elgar.

Gonzalez, R. and Wu, G. (1999). 'On the shape of the probability weighting function', *Cognitive Psychology*, vol. 38, pp. 129–66.

Harless, D. W. and Camerer, C. F. (1994). 'The predictive utility of generalized expected utility theories', *Econometrica*, vol. 62, pp. 1251–89.

Harrison, G. W. (1994). 'Expected utility theory and the experimentalists', *Empirical Economics*, vol. 19(2), pp. 223–53.

© The Author(s). Journal compilation © Royal Economic Society 2009

Harrison, G. W., Johnson, E., McInnes, M. M. and Rutström, E. E. (2005). 'Risk aversion and incentive effects: comment', *American Economic Review*, vol. 95(3), pp. 897–901.
Harrison, G. W., Lau, M. I. and Rutström, E. E. (2007). 'Estimating risk attitudes in Denmark: a field experiment', *Scandinavian Journal of Economics*, vol. 109(2), pp. 341–68.
Harrison, G. W., Lau, M. I., Rutström, E. E. and Sullivan, M. B. (2005). 'Eliciting risk and time preferences using field experiments: some methodological issues', in (J. Carpenter, G.W. Harrison and J.A. List, eds.), *Field Experiments in Economics, Research in Experimental Economics*, vol. 10, pp. 125–218, Greenwich, CT: JAI Press.
Harrison, G. W. and List, J. A. (2004). 'Field experiments', *Journal of Economic Literature*, vol. 42(4), pp. 1013–59.
Harrison, G. W., List, J. A. and Towe, C. (2007). 'Naturally occurring preferences and exogenous laboratory experiments: a case study of risk aversion', *Econometrica*, vol. 75(2), pp. 433–58.
Harrison, G. W. and Rutström, E. E. (2008). 'Risk aversion in the laboratory', in (J. Cox and G.W. Harrison, eds.), *Risk Aversion in Experiments, Research in Experimental Economics*, vol. 12, pp. 41–196, Bingley: Emerald.
Harrison, G. W. and Rutström, E. E. (forthcoming). 'Representative agents in lottery choice experiments: one wedding and a decent funeral', *Experimental Economics*.
Hertwig, R., Barron, G., Weber, E. U. and Erev, I. (2004). 'Decisions from experience and the effect of rare events in risky choice', *Psychological Science*, vol. 15(8), pp. 534–9.
Heston, A., Summers, R. and Aten, B. (2002). *Penn World Table Version 6.1*, Center for International Comparisons at the University of Pennsylvania (CICUP), data accessed at http://pwt.econ.upenn.edu/php_site/pwt61_form.php.
Hey, J. D. and Orme, C. (1994). 'Investigating generalizations of expected utility theory using experimental data', *Econometrica*, vol. 62(6), pp. 1291–326.
Holt, C. A. and Laury, S. K. (2002). 'Risk aversion and incentive effects', *American Economic Review*, vol. 92(5), pp. 1644–55.
Humphrey, S. J. (2000). 'The common consequence effect: testing a unified explanation of recent mixed evidence', *Journal of Economic Behaviour & Organization*, vol. 41, pp. 239–62.
Humphrey, S. J. and Verschoor, A. (2004*a*). 'Decision-making under risk among small farmers in East Uganda', *Journal of African Economies*, vol. 13(1), pp. 44–101.
Humphrey, S. J. and Verschoor, A. (2004*b*). 'The probability weighting function: experimental evidence from Uganda, India and Ethiopia', *Economics Letters*, vol. 84, pp. 419–25.
Kahneman, D. and Tversky, A. (1979). 'Prospect theory: an analysis of decision under risk', *Econometrica*, vol. 47, pp. 263–91.
Levy, H. and Levy, M. (2002). 'Arrow-Pratt risk aversion, risk premium and decision weights', *Journal of Risk & Uncertainty*, vol. 25(3), pp. 265–90.
Liang, K-Y. and Zeger, S.L. (1986). 'Longitudinal data analysis using generalized linear models', *Biometrika*, vol. 73, pp. 13–22.
List, J. A. (2004). 'Neoclassical theory versus prospect theory: evidence from the marketplace', *Econometrica*, vol. 72(2), pp. 615–25.
Loomes, G. and Sugden, R. (1995). 'Incorporating a stochastic element into decision theories', *European Economic Review*, vol. 39, pp. 641–8.
Mosley, P. and Verschoor, A. (2005). 'Risk attitudes and the "vicious circle of poverty" ', *European Journal of Development Research*, vol. 17(1), pp. 59–88.
Prelec, D. (1998). 'The probability weighting function', *Econometrica*, vol. 6, pp. 497–527.
Quiggin, J. (1982). 'A theory of anticipated utility', *Journal of Economic Behaviour & Organization*, vol. 3(4), pp. 323–43.
Rieger, M. O. and Wang, M. (2006). 'Cumulative prospect theory and the St. Petersburg paradox', *Economic Theory*, vol. 28, pp. 665–79.
Rogers, W. H. (1993). 'Regression standard errors in clustered samples', *Stata Technical Bulletin*, vol. 13, pp. 19–23.
Rubinstein, A. (2002). 'Comments on the risk and time preferences in economics', mimeo, Department of Economics, Princeton University.
Starmer, C. (1992). 'Testing new theories of choice under uncertainty using the common consequence effect', *Review of Economic Studies*, vol. 59, pp. 813–30.
Starmer, C. (2000). 'Developments in non-expected utility theory developments in non-expected utility theory: the hunt for a descriptive theory of choice under risk', *Journal of Economic Literature*, vol. 38 (June) pp. 332–82.
Tversky, A. and Kahneman, D. (1992). 'Advances in prospect theory: cumulative representation of uncertainty', *Journal of Risk and Uncertainty*, vol. 5, pp. 297–323.
Williams, R. L. (2000). 'A note on robust variance estimation for cluster-correlated data', *Biometrics*, vol. 56 (June) pp. 645–6.
Wooldridge, J. (2003). 'Cluster-sample methods in applied econometrics', *American Economic Review (Papers & Proceedings)*, vol. 93 (May), pp. 133–8.
Yaari, M. E. (1987). 'The dual theory of choice under risk', *Econometrica*, vol. 55(1), pp. 95–115.

© The Author(s). Journal compilation © Royal Economic Society 2009